

Contamination cannot explain the lack of large-scale power in the cosmic microwave background radiation

Emory F. Bunn* and Austin Bourdon
Physics Department, University of Richmond, Richmond, VA 23173

Several anomalies appear to be present in the large-angle cosmic microwave background (CMB) anisotropy maps of WMAP. One of these is a lack of large-scale power. Because the data otherwise match standard models extremely well, it is natural to consider perturbations of the standard model as possible explanations. We show that, as long as the source of the perturbation is statistically independent of the source of the primary CMB anisotropy, no such model can explain this large-scale power deficit. On the contrary, any such perturbation always *reduces* the probability of obtaining any given low value of large-scale power. We rigorously prove this result when the lack of large-scale power is quantified with a quadratic statistic, such as the quadrupole moment. When a statistic based on the integrated square of the correlation function is used instead, we present strong numerical evidence in support of the result. The result applies to models in which the geometry of spacetime is perturbed (e.g., an ellipsoidal Universe) as well as explanations involving local contaminants, undiagnosed foregrounds, or systematic errors. Because the large-scale power deficit is arguably the most significant of the observed anomalies, explanations that worsen this discrepancy should be regarded with great skepticism, even if they help in explaining other anomalies such as multipole alignments.

PACS numbers: 98.80.-k, 98.70.Vc, 98.80.Es, 95.85.Bh

I. INTRODUCTION

Observations of cosmic microwave background (CMB) anisotropy, particularly the data from WMAP [1, 2, 3, 4], have revolutionized cosmology. These observations are a major contributor to the emergence of a cosmological “standard model” of a Universe dominated by dark energy and cold dark matter, with a nearly scale-invariant spectrum of Gaussian adiabatic perturbations [5, 6]. The overall consistency of the CMB data with this model is quite remarkable, but there appear to be some anomalies on the largest angular scales, such as a lack of large-scale power [2, 7, 8], alignment of low-order multipoles [8, 9, 10, 11], and hemispheric asymmetries [12, 13].

The significance of and explanations for these puzzles are hotly debated. In particular, it is difficult to know how to interpret a posteriori statistical significances: when a statistic is invented to quantify an anomaly that has already been noticed, the low p -values for that statistic cannot be taken at face value. Nonetheless, the number and nature of the anomalies (in particular, the fact that several seem to pick out the same directions on the sky) seem to suggest that there may be something to explain in the data. In this paper, we will tentatively assume that there is a need for an explanation and consider what that explanation might be.

Since the standard model is in general highly consistent with the CMB and a wide variety of other observations, it is natural to look for explanations of these puzzles that consist of perturbations added onto the standard model. Such explanations can be based on nonstandard cosmolo-

gies, such as ellipsoidal models [14], large-scale magnetic fields [15], and theories based on Bianchi VIIh spacetimes with rotation and shear [16, 17]. They can also involve phenomena on much smaller scales (e.g., [18, 19, 20, 21]), perhaps even within the Solar System [22]. Any undiagnosed foreground contaminant would fall into the class of explanations we consider, as would many systematic errors.

All of these models can be described by assuming that the observed CMB sky is the sum of two terms:

$$T_{\text{obs}}(\mathbf{r}) = T_0(\mathbf{r}) + T_c(\mathbf{r}), \quad (1)$$

where T_0 is a Gaussian CMB sky with a power spectrum given by the standard model and T_c is a contaminant. The contaminant can be a fixed function of sky position \mathbf{r} or a realization of a random process. In the latter case, we assume nothing about the statistics of this process except that it is independent of the Gaussian random process that produced T_0 . We wish to consider the possibility that such a model can explain some or all of the large-angle anomalies.

In this article, we will present strong evidence that on the contrary all such models actually exacerbate one of the anomalies, namely the observed lack of power in the large-angular-scale CMB anisotropy. This anomaly is formally highly statistically significant, and as we will argue below it is one for which the problems of a posteriori statistics are not particularly severe. It is therefore arguably the most in need of explanation of all of the large-angle CMB puzzles. We conclude, therefore, that this entire category of possible explanations should be regarded with great skepticism. In particular, the absence of large-scale power in the WMAP data is in fact a strong argument *against* the existence of undiagnosed foreground contamination, as well as systematic errors

*Electronic address: ebunn@richmond.edu

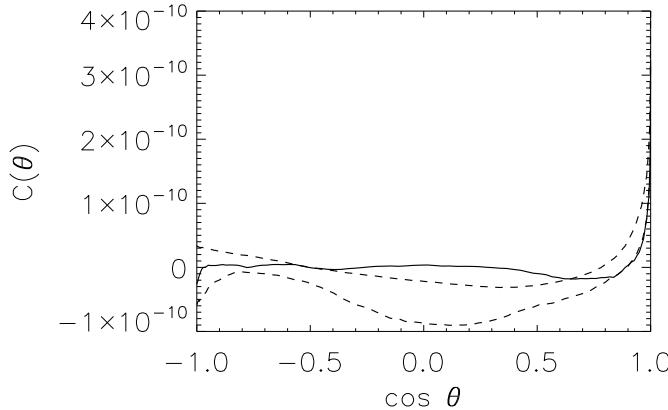


FIG. 1: The two-point correlation function for the WMAP data. The solid curve shows the correlation function for the three-year WMAP internal linear combination (ILC) data at resolution $N_{\text{side}} = 32$, computed with the SpICE software [24]. The dashed curves show the 95% confidence range in a set of simulations. The simulations virtually never produce correlation functions that are as close to zero at large angles as the real data.

that would produce an additive contaminant to the observed sky maps.

We can quantify the lack of power in the large-angle CMB by considering either the power in low-order multipoles (especially the quadrupole) or a statistic based on the two-point angular correlation function (see Fig. 1). In either case, the p -values (that is, the probabilities of getting as low a value of the chosen statistic as the one in the actual data) are low; in fact, for some choices of statistic, they are less than 0.1% [7] (but see [23] for a contrasting analysis). By definition, for an alternative theory to explain this anomaly, it would have to generate larger p -values. We will show in this paper that all proposed models of the form described above in fact *reduce* the p -values based on these statistics. Therefore, although such models might alleviate some of the other large-angle anomalies, they worsen this one.

At one level, this is not surprising. For the models considered here, in which the observations are the sum of two statistically independent terms, the observed power spectrum is simply the sum of the standard-model spectrum and the spectrum of the contaminant. Addition of the contaminant therefore biases all multipoles up, including the quadrupole. This is merely a statement about mean-square values, however, and does not tell us about the probability distribution of the multipoles. It is logically possible that a (non-Gaussian) contaminant, even as it biases the mean-square quadrupole up, widens the probability distribution for the quadrupole in such a way as to enhance the probability of getting low values. Indeed, any proposal to explain the lack of large-scale power through a perturbation to the standard model must be

proposing such an effect, since this is what it would mean to “explain” the discrepancy.

For example, suggestions have been made that the low quadrupole might be explained by an extended local foreground [20], by dust-filled local voids¹ [18, 19], or by an “ellipsoidal” universe that expands at different rates in different directions [14]. Each such explanation assumes that a chance anticorrelation between the contaminant and the intrinsic CMB anisotropy has occurred. In order for this to count as an explanation, however, such an anticorrelation must be sufficiently probable that it raises the probability of finding the observed lack of power. Although this is a logical possibility, we will argue below that it in fact never occurs, whether the lack of power is quantified via the quadrupole moment or the correlation function. For some specific cases, such as the quadrupole moment in an ellipsoidal universe, previous work [25] has already established this; in this paper we prove it in general. In summary, such models cannot explain the lack of large-scale power, and in fact always “anti-explain” it by reducing the already-low probability.

Section II proves this general result in the case where the lack of power is quantified via a quadratic estimator such as the mean-square quadrupole moment. Section III presents strong numerical evidence that the result is also true in the case of a statistic based on the two-point correlation function. Section IV contains a brief discussion of the results, and an appendix proves a key mathematical result needed in section II.

II. QUADRATIC POWER ESTIMATORS

As noted above, the observed lack of large-scale power in the CMB can be quantified in different ways. The simplest, going all the way back to the COBE observations [26, 27], is to compute an estimator of the quadrupole power $C_2 = \langle |a_{2m}|^2 \rangle$, where a_{lm} is a coefficient in a spherical harmonic expansion. Quadrupole estimators applied to the WMAP data are lower than theoretical predictions, although due to the large cosmic variance, the significance of this anomaly is only $\sim 5\%$ [8], which is weaker than the correlation function statistic described in the next section. Nonetheless, because the quadrupole is one of the simplest and most natural ways to quantify large-scale power, we consider it in detail in this section. In particular, we will demonstrate that any statistically independent contaminant exacerbates the problem of an anomalously low quadrupole.

¹ A suggestion is made in the cited work that the hypothesis that the contaminant is uncorrelated with the primary signal may not apply. If this is true, then the arguments in the present paper would not apply to this model. It is not clear to us that a strong correlation of the proposed form exists in the model under consideration, and as far as we know no detailed calculation of this effect has been performed.

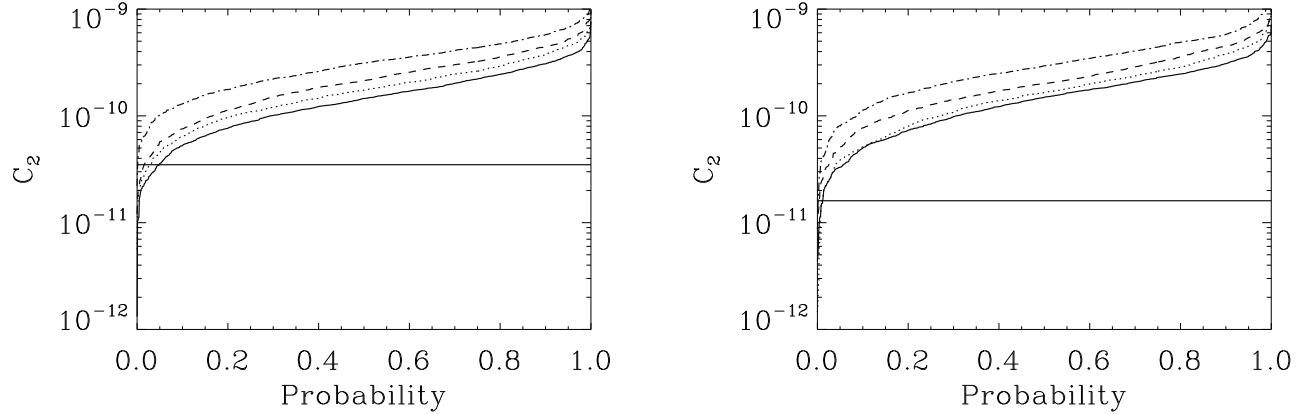


FIG. 2: Cumulative probability distributions for the quadrupole power C_2 in an ellipsoidal Universe, in dimensionless $\Delta T/T$ units. In the left panel, data from the entire sky were used, and in the right panel the WMAP Kp0 cut was applied. The solid curve shows a standard LCDM model with no eccentricity. From bottom to top, the other three curves correspond to eccentricities 5×10^{-3} , 6.2×10^{-3} , 7.4×10^{-3} . The horizontal line shows the quadrupole found in the actual data.

The quadrupole power is a positive definite quadratic function q^2 of the data. As noted in the previous section, a contaminant always causes an upward bias in the expectation value of such a statistic; to be precise, the expectation value is $\langle q^2 \rangle = \langle q_0^2 \rangle + \langle q_c^2 \rangle$, where the two terms on the right are the expectation values due to the contributors T_0, T_c . As noted in the previous section, however, this statement is not sufficient to justify the claim that adding a contaminant always exacerbates the problem of a low quadrupole. We need to show that the probability of getting a low quadrupole is always reduced by adding a contaminant – that is, for any given value \hat{q}^2 , the probability that the observed value is less than \hat{q}^2 is always lower with a contaminant than without.

Let the vector \mathbf{y} represent a list of data points that we will use to estimate the large-angle power in the CMB, for example, the pixelized temperature values in the WMAP data. Let q^2 be a positive definite quadratic function of the data (possibly with some noise bias removed):

$$q^2(\mathbf{y}) = \mathbf{y} \cdot \mathbf{A} \cdot \mathbf{y} - b. \quad (2)$$

Here \mathbf{A} is a symmetric nonnegative definite matrix, and the noise bias b is a constant.

We want to compare the null hypothesis, that \mathbf{y} contains only intrinsic CMB anisotropy and noise, with the hypothesis that there is an additional statistically independent contaminant. We can express these possibilities by writing

$$\mathbf{y} = \mathbf{x} + \mathbf{c}, \quad (3)$$

where \mathbf{x} is the “uncontaminated” data (including noise) and \mathbf{c} represents a hypothetical contaminant. We assume that \mathbf{x} is drawn from a multivariate Gaussian distribution:

$$f_x(\mathbf{x}) \propto \exp\left(-\frac{1}{2}\mathbf{x} \cdot \mathbf{M} \cdot \mathbf{x}\right) \quad (4)$$

for some symmetric positive definite matrix \mathbf{M} . For the null hypothesis, we set $\mathbf{c} = 0$. When considering contamination, we assume \mathbf{c} is a random variable with some probability density f_c . (This formulation includes the possibility that \mathbf{c} is a fixed contaminant – that is, f_c is allowed to be a delta function.) No assumption is made about f_c other than independence of \mathbf{x} and \mathbf{c} , which means that the joint probability density factors:

$$f(\mathbf{x}, \mathbf{c}) = f_x(\mathbf{x})f_c(\mathbf{c}). \quad (5)$$

Let $\hat{\mathbf{y}}$ be the data actually measured, and let $\hat{q}^2 = q^2(\hat{\mathbf{y}})$ stand for the power estimate obtained from it. Let P_c stand for the probability of getting a value of q^2 as low as the true value, assuming a fixed value for the contaminant \mathbf{c} :

$$P_c = \Pr[q^2(\mathbf{y}) < \hat{q}^2 \mid \mathbf{c}] = \int_{(\mathbf{x}+\mathbf{c}) \in V} d\mathbf{x} f_x(\mathbf{x}), \quad (6)$$

where the volume V is the ellipsoid consisting of all \mathbf{y} with $\mathbf{y} \cdot \mathbf{A} \cdot \mathbf{y} < \hat{q}^2 + b$.

Note that P_c is an integral over an ellipsoid centered at $\mathbf{x} = -\mathbf{c}$. Since the integrand peaks at the origin, we would expect P_c to be maximized when the ellipsoid’s center is placed at the origin. To be specific, we expect that

$$P_c \leq P_0. \quad (7)$$

This expectation is indeed correct; a proof of it may be found in the Appendix.

This means that adding any fixed contaminant \mathbf{c} always reduces the probability of getting a low q^2 . As a consequence, even if \mathbf{c} is not fixed but is generated by some random process, the probability is still lower than in the case $\mathbf{c} = 0$. Formally, we can write

$$\Pr[q^2(\mathbf{y}) < \hat{q}^2] = \int d\mathbf{c} \Pr[q^2(\mathbf{y}) < \hat{q}^2 \mid \mathbf{c}]f_c(\mathbf{c}). \quad (8)$$

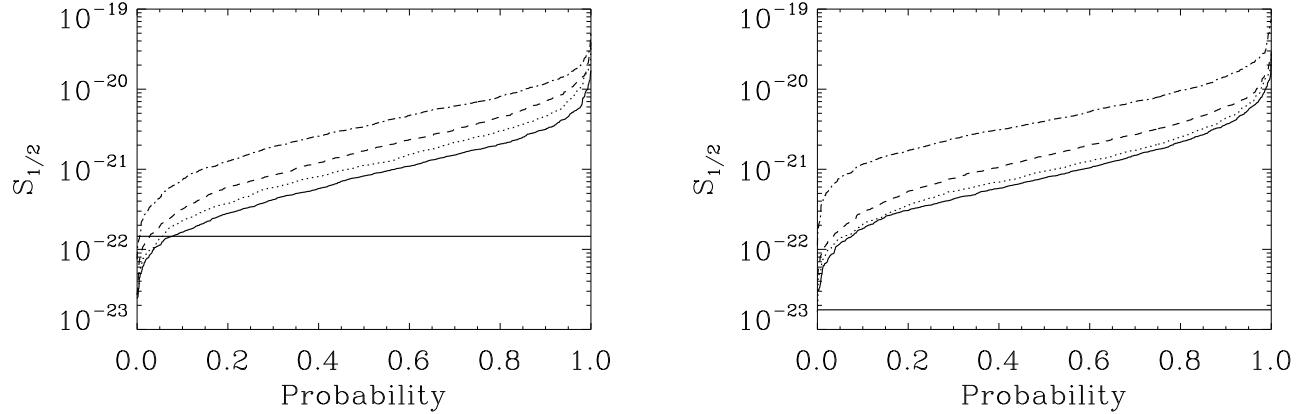


FIG. 3: Cumulative probability distributions for the statistic $S_{1/2}$. As in Figure 2, the left panel is for full-sky data, while the right panel is for the Kp0 cut. Curves are as in the previous figure.

Using inequality (7),

$$\Pr[q^2(\mathbf{y}) < \hat{q}^2] \leq P_0 \int f_c(\mathbf{c}) d\mathbf{c} = P_0. \quad (9)$$

This inequality is the central result of this section. It means that, if the power is anomalously low under the assumption of no contamination, then introducing a contaminant can only make the problem worse.

Figure 2 illustrates this conclusion for the case of an ellipsoidal Universe. The figure shows the cumulative probability distribution of the quadrupole power C_2 , based on 1000 simulations of the CMB sky. The simulations were performed using HEALPix [28] with $N_{\text{side}} = 32$. The solid curve shows the distribution for a Gaussian CMB with the power spectrum given by the best-fit LCDM model from the three-year WMAP data [29]. From bottom to top, the other three curves show models with the same power spectrum but with eccentricities $5 \times 10^{-3}, 6.2 \times 10^{-3}, 7.4 \times 10^{-3}$. According to the analysis of ref. [14], eccentricities in this range provide a better fit to the CMB quadrupole than the standard model; however, as Gruppuso [25] has pointed out, the calculations in ref. [14] do not properly account for all possible relative orientations of the ellipticity axis and the intrinsic CMB anisotropy and hence overestimate the goodness of fit of the ellipsoidal models. The horizontal line indicates the value found in the actual WMAP data (specifically, the three-year internal linear combination data, downgraded to $N_{\text{side}} = 32$). The curves in the left panel were computed using the entire sky, while the right curves were computed using the WMAP Kp0 cut [3].

The figure illustrates that the probability of getting a quadrupole value below any given cutoff strictly decreases as the size of the perturbation increases. As predicted by inequality (9), the way to get the highest probability is to have no perturbation at all. In particular, for the no-cut data, the probability of getting a value as small as the actual data is $\sim 5\%$ in the standard

model and drops to $\sim 3\%, 1.5\%, 0.2\%$ as the ellipticity increases. When the Kp0 mask is applied, the probabilities are lower in all cases than in the full-sky case, but the same decrease in probability is observed. These conclusions are consistent with those of ref. [25], but we have established the conclusion for a much broader category of theories, not just this specific case.

III. CORRELATION FUNCTION

The low quadrupole does not have particularly high statistical significance, largely because of the high level of cosmic variance in the quadrupole. The two-point angular correlation function provides a much more significant indication that there is an anomalous lack of large-scale power in the WMAP data. In particular, the integrated square of the correlation function,

$$S_{1/2} = \int_{-1}^{1/2} [C(\theta)]^2 d\cos\theta, \quad (10)$$

which was first introduced in the analysis of the 1-year WMAP data [30], is extremely low in the WMAP data in comparison with theoretical estimates, with p -value of order 0.1% [7]. Here $C(\theta)$ is the two-point correlation function, that is, the average of all pairs of pixels with angular separation θ . We wish to examine whether adding a perturbation to the standard model can solve this problem (that is, raise the probability of getting the observed low value of $S_{1/2}$).

Since the statistic $S_{1/2}$ is quartic, not quadratic, in the data, the argument of the previous section does not apply to it. However, it is extremely plausible to suppose that a similar conclusion should hold, since any model with a high probability of producing low values of this statistic would presumably produce low values of the lower-order multipoles, and since any contaminant reduces the

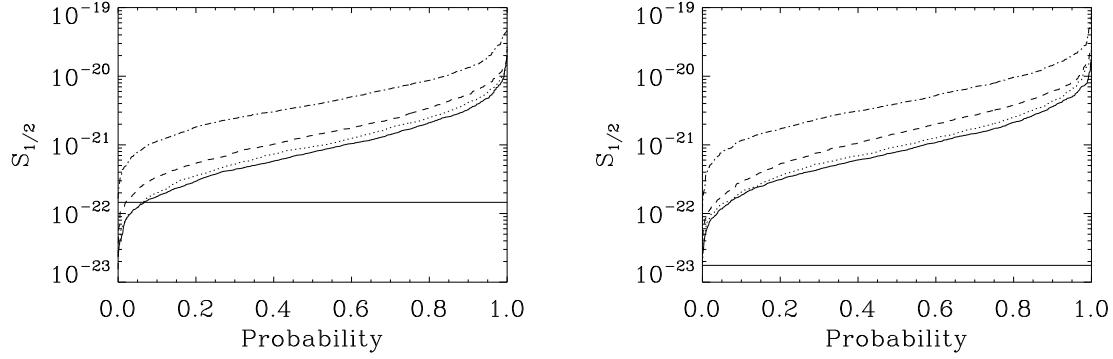


FIG. 4: Cumulative probability distributions for the statistic $S_{1/2}$ for Bianchi VIIh (rotating Universe) models. The values of the shear are 0 (solid), 2.4×10^{-10} (dotted), 5×10^{-10} (dashed), and 1×10^{-9} (dot-dashed). As in the previous figures, no-cut probabilities are shown on the left, and Kp0-cut probabilities are shown on the right.

probability of such low multipoles.

We can of course test this conjecture numerically for any particular model. For example, Figure 3 shows the results of simulations precisely like those shown in Figure 2, but with the statistic $S_{1/2}$ used in place of the quadrupole. The SpICE software [24] was used to compute the correlation functions. Figure 4 shows the results of similar calculations, for the case of a model in which the spacetime geometry is that of a rotating Bianchi VIIh model [17]. We have also performed computations for models in which the contaminant consists of circular hot and cold spots of varying amplitudes and radii, to simulate the effects of local voids or similar features. In all of these cases, the addition of a contaminant does not solve the problem of the lack of large scale power; in fact, it worsens it.

Rather than examining theories one at a time, it would clearly be better to have a general argument that applied to a broad class of theories. In the rest of this section, we provide such an argument.

Suppose that the value of $S_{1/2}$ for the actual data is $\hat{S}_{1/2}$. Let V be the volume in the data space that yields values of the statistic this low:

$$V = \{\mathbf{y} \mid S_{1/2}(\mathbf{y}) < \hat{S}_{1/2}\}. \quad (11)$$

Then, assuming a contaminant given by a fixed vector \mathbf{c} , the probability of getting such a low value of the statistic is

$$P(\mathbf{c}) = \int_{(\mathbf{x}+\mathbf{c}) \in V} f_x(\mathbf{x}) d\mathbf{x} = \int_{\mathbf{y} \in V} f_x(\mathbf{y} - \mathbf{c}) d\mathbf{y}. \quad (12)$$

We want to know whether there are any vectors \mathbf{c} such that $P(\mathbf{c}) > P(0)$, or in other words whether P has a global maximum at $\mathbf{c} = 0$. It is straightforward to check that $\nabla P(0) = 0$. We next consider whether the point $\mathbf{c} = 0$ is a maximum, a minimum, or a saddle point. If we find that it is a maximum, then the addition of any small contaminant worsens the problem we are trying to solve.

To answer this question, we naturally consider the matrix of second derivatives:

$$H_{jk} = - \left. \frac{\partial P}{\partial c_j \partial c_k} \right|_{\mathbf{c}=0}. \quad (13)$$

Then P has a local maximum at the origin if and only if \mathbf{H} is positive definite. Moreover, if \mathbf{H} is not positive definite, then the eigenvectors corresponding to negative eigenvalues yield the directions in data space (i.e., particular forms for the contaminant \mathbf{c}) that alleviate the problem of low $S_{1/2}$.

To calculate these derivatives, it is convenient to transform the data to a basis that diagonalizes the covariance matrix in the Gaussian probability density f_x . The most natural way to accomplish this is to work in the spherical harmonic basis, in which case each data point is a coefficient a_{lm} . We can normalize each data point according to the power spectrum, setting $x_j = a_{lm}/C_l^{1/2}$, where the index j runs over all pairs lm . In this case the covariance matrix is simply the identity matrix, and the second derivative matrix elements can be written

$$H_{jk} = - \int_V d\mathbf{x} f_x(\mathbf{x}) (x_j x_k - \delta_{jk}). \quad (14)$$

This integral over the many-dimensional data space can most easily be estimated by Monte Carlo integration. To be specific, we draw vectors \mathbf{x} from the appropriate Gaussian distribution, calculate the corresponding values of $S_{1/2}$, and use the results to throw away all vectors that lie outside of V . For all the rest, we average together the quantities $(x_j x_k - \delta_{jk})$.

In performing this Monte Carlo integration, we consider HealPIX maps with $N_{\text{side}} = 32$ and the same power spectrum as in the previous section. We apply Gaussian smoothing with a 20° FWHM beam to the simulated maps. This amount of smoothing results in significant suppression (by more than e^{-1}) of spherical harmonics coefficients $l \geq 10$. Without significant smoothing, fluctuations in high- l modes cause significant error in the

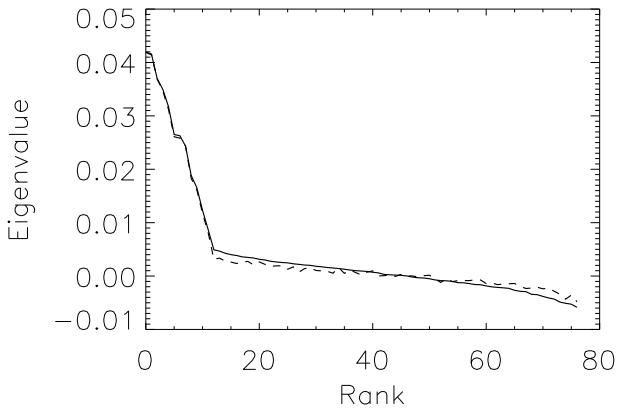


FIG. 5: Eigenvalues of the second derivative matrix \mathbf{H} . Two independent calculations of the matrix were performed, each based on 80 000 simulations using the Kp0 cut. The matrices were truncated to include only multipoles $l = 2$ through 8. The solid curve shows the eigenvalues computed from one matrix, sorted from largest to smallest. The dashed curve shows the quantities $\mathbf{v} \cdot \mathbf{H} \cdot \mathbf{v}$, where \mathbf{v} are the eigenvectors computed from the first matrix and \mathbf{H} is the second matrix. The difference between the two curves gives an indication of the numerical error in the Monte Carlo integration.

Monte Carlo calculation even at low l . The problem of anomalously low $S_{1/2}$ persists at about the same significance (p -values $\simeq 0.1\%$) even with such smoothing, so this smoothing does not weaken our ability to draw conclusions about possible explanations for the anomaly.

Figure 5 shows the eigenvalues of the matrix resulting from this Monte Carlo integration, using the Kp0 mask. The results look similar when data from the full sky are used. The matrix used to compute the eigenvalues was based on 80 000 simulations lying within the volume V . Modes up to $l = 8$ were used to compute the eigenvalues shown in the figure, although modes up to $l = 64$ (far above the beam scale) were used in the simulations. To test the numerical stability of the results, we used a second set of 80 000 simulations to recompute the matrix \mathbf{H} . We then calculated $\mathbf{v} \cdot \mathbf{H} \cdot \mathbf{v}$ for each eigenvector \mathbf{v} . The results are shown in the dashed curve. In the absence of numerical error, the two curves would be identical.

Although there is some numerical error due to the Monte Carlo integration, it appears that the matrix is not positive definite. We wish to examine the eigenvectors corresponding to the most negative eigenvalues, since these describe particular contaminants that might solve the problem of a lack of large-scale power. Figure 6 shows the particular pattern on the sky corresponding to the most negative eigenvalue. Most of the power in this contaminant is found in multipole $l = 5$, as is the power in all of the most negative eigenvectors. To test the robustness of this pattern, we computed the eigenvectors retaining varying numbers of modes in the matrix

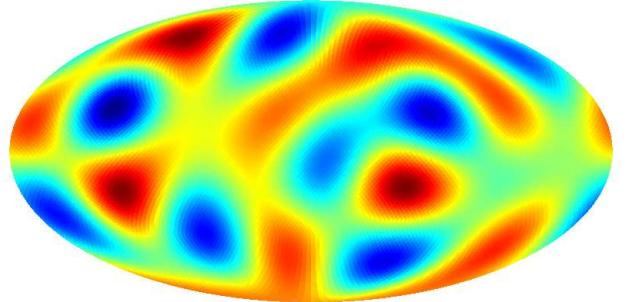


FIG. 6: The sky pattern corresponding to the most negative eigenvalue of \mathbf{H} .

\mathbf{H} , ranging from $l_{\max} = 5$ to 15, and also using varying subsets of the Monte Carlos to compute the matrix. The results are quite consistent, with the most negative eigenvectors always having most of their power at $l = 5$ and looking quite similar to Fig. 6.

The existence of these negative eigenvalues seems to contradict our assertion that no contaminant can explain the low value of $S_{1/2}$: modes such as the one shown in Fig. 6, by construction, raise the probability of getting a low value when added to the data. However, when we assess the amount of improvement that these modes can provide, we find it to be negligible. Consider a model in which we add a contaminant of the form shown in Fig. 6 with some amplitude α to the standard model. The results of this section have shown that the probability of getting a low $S_{1/2}$ is an increasing function of α at low α . However, because the eigenvalue is fairly small, the increase might be expected to be slight. Furthermore, for sufficiently large value of α , the probability must start to decrease again.

Fig. 7 shows that this is indeed the case, and furthermore that no choice of α leads to a significant increase in the probability of getting a value as low as the actual data. This probability remains virtually unchanged at $\sim 10^{-3}$ for small α and then decreases dramatically for larger α . Since all of the eigenvectors corresponding to significantly negative eigenvalues of \mathbf{H} give patterns quite similar to this one, we can conclude with confidence that no such pattern can significantly alleviate the problem of low $S_{1/2}$.

IV. CONCLUSIONS

We have considered a broad class of cosmological models, obtained by adding a contaminant to the standard best-fit inflation-based model. The only assumption we have made about the contaminant is that it is statistically independent of the cosmological signal. We have argued that all such models exacerbate rather than alleviating the lack of large-scale power in the WMAP data. We

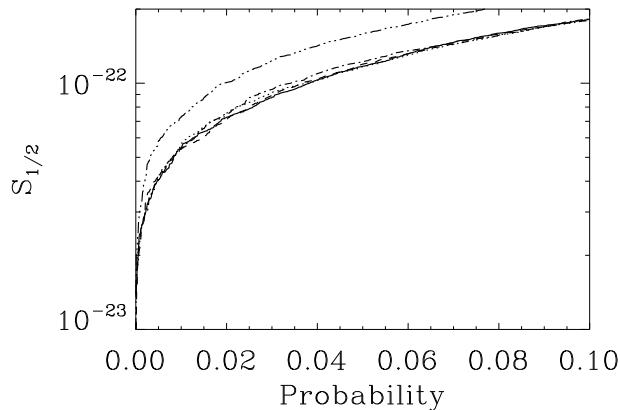


FIG. 7: Cumulative probability distributions for models in which a fixed contaminant of the form shown in Fig. 6 is added to the sky. This pattern corresponds to the most negative eigenvalue of the matrix \mathbf{H} and so might be expected to increase the probability of finding a low value of $S_{1/2}$. The curves shown in the figure are for varying amplitudes of the contaminant, with root-mean-square pixel values of 0 (solid), $2 \mu\text{K}$ (dotted), $4 \mu\text{K}$ (dashed), $8 \mu\text{K}$ (dot-dashed). No value of the amplitude causes the probability of getting values as low as those in the real data to increase noticeably.

have proven this result to be true when the lack of power is quantified by the quadrupole moment and have presented strong numerical evidence in support of it when the two-point correlation function is used. Since the latter in particular is discrepant at a highly significant level already, any theory that worsens this discrepancy should be regarded with great skepticism.

In addition to exotic cosmologies such as models with a global ellipsoidal anisotropy, the class of models considered herein includes more mundane possibilities such as undiagnosed foregrounds and many systematic errors. In particular, since several of the observed anomalies seem to “pick out” the ecliptic plane as a preferred direction, some attention has focused on a local foreground as a possible explanation. The calculations presented here argue against such models.

In any particular model with a contaminant, of course, it is possible that a chance cancellation between the contaminant and the intrinsic CMB anisotropy can occur, leading to the observed lack of large-scale power. What we have shown is that such a chance cancellation is always unlikely, and in particular that it is always more unlikely than the lack of large-scale power occurring based on cosmic variance alone, without a contaminant.

The question of how seriously to take the various large-angle CMB anomalies, including the lack of large-scale power as well as the various other puzzles, has been much debated [31]. In particular, because they are all based on a posteriori statistics (*i.e.*, on statistical significances calculated after the anomalies had already been noted),

the quoted significances cannot be taken at face value. Arguably, however, the large-scale power deficit suffers less from this problem of a posteriori statistics. After all, for virtually the entire existence of the field of CMB anisotropy studies, the two-point correlation function has been regarded as one of the most natural statistics to use in quantifying the level of structure in CMB maps as a function of angular scale. For instance, upper limits on CMB anisotropy in the pre-COBE era were usually presented as limits on the correlation function. Although the particular statistic $S_{1/2}$ is an a posteriori invention, it merely quantifies the mean-square level of this function, which was already regarded *a priori* as a natural function to compute. Although one can certainly dispute the extent to which the significance of the lack of large-angle correlations is an artifact of the particular choice of statistic (for instance, see [32], who do not use the $S_{1/2}$ statistic and find less significant discrepancies), nonetheless we believe that, of all the observed anomalies, the large-scale power deficit is one of the most in need of explanation.

Anomalies in a data set naturally prompt thoughts of systematic errors or contaminants in the data. Perhaps counterintuitively, this particular anomaly provides a strong argument against such possibilities. In particular, a foreground that was not removed from the data (due to having a spectrum indistinguishable from the CMB, for example) would fall precisely into the category considered herein. Note, however, that if the foreground removal procedure itself removes part of the cosmological signal, the resulting error would not fall into the category considered herein. In particular, the ILC method does project out some of the intrinsic CMB signal and so in principle does reduce the amount of large-scale power. This effect is calculable and has been found to be negligible, however.

There are of course a wide variety of possible explanations for the anomalies that do not fall into the category considered here. For example, simply modifying the primordial power spectrum at large scales naturally alleviates the problem of a lack of large-scale CMB power (e.g., [33, 34]). Some models with nontrivial topology (e.g., [35, 36, 37, 38, 39]) also have this effect, although such models have other problems [40]. The framework of spontaneous isotropy breaking [41] also provides a class of models that are not based on simply adding a perturbation to the standard cosmology. Models such as these (and many others) may provide an explanation for the puzzles in the large-angle CMB.

Acknowledgments

This work was supported by National Science Foundation Award AST-0507395, by the Research Corporation, and by the Virginia Foundation for Independent Colleges. We acknowledge use of the HEALPix [28] and SpICE [24] software.

APPENDIX: PROOF OF EQUATION (7)

We first express equation (6) in terms of the integration variable $\mathbf{y} = \mathbf{x} + \mathbf{c}$:

$$P_{\mathbf{c}} = \int_{\mathbf{y} \in V} d\mathbf{y} f_x(\mathbf{y} - \mathbf{c}). \quad (15)$$

We next apply a linear coordinate transformation that maps the ellipsoid V onto the unit sphere. To be specific, we find a matrix \mathbf{L} such that $\mathbf{A} = \mathbf{L} \cdot \mathbf{L}^T$ (e.g., by Cholesky decomposition). We define $\mathbf{y}' = (\mathbf{L}^T \cdot \mathbf{y}) / \sqrt{q_0^2 + b}$ and \mathbf{c}' similarly. Then

$$P_{\mathbf{c}} \propto \int_{|\mathbf{y}'|^2 \leq 1} d\mathbf{y}' f_{x'}(\mathbf{y}' - \mathbf{c}') \quad (16)$$

Here $f_{x'}$ is a multivariate Gaussian probability density with a new inverse covariance matrix \mathbf{M}' , and the proportionality constant is determined by the Jacobian of the coordinate transformation. For convenience, we now make yet another coordinate transformation: we apply a rotation that diagonalizes \mathbf{M}' . The result is

$$P_{\mathbf{c}} \propto \int_{|\mathbf{y}''|^2 \leq 1} d\mathbf{y}'' \exp \left(-\sum \frac{(y''_i - c''_i)^2}{2\sigma_i^2} \right). \quad (17)$$

For the remainder of this section we drop the double primes.

We now show that $P_{\mathbf{c}}$ has a maximum at $\mathbf{c} = 0$. Differentiate the above expression for $P_{\mathbf{c}}$ with respect to c_1 :

$$\frac{\partial P_{\mathbf{c}}}{\partial c_1} \propto \int_{|\mathbf{y}| \leq 1} dy \exp \left(-\sum \frac{(y_i - c_i)^2}{2\sigma_i^2} \right) \frac{y_1 - c_1}{\sigma_1^2} \quad (18)$$

$$= \int dy_2 \cdots dy_n \exp \left(-\sum_{i=2}^n \frac{(y_i - c_i)^2}{2\sigma_i^2} \right) \times \int_{-Y_1}^{Y_1} \exp \left(-\frac{(y_1 - c_1)^2}{2\sigma_1^2} \right) \frac{y_1 - c_1}{\sigma_1^2} dy_1, \quad (19)$$

where $Y_1 = (1 - \sum_{i=2}^n y_i^2)^{1/2}$. Performing the y_1 integral yields

$$\frac{\partial P_{\mathbf{c}}}{\partial c_1} \propto \int dy_2 \cdots dy_n \exp \left(-\sum_{i=2}^n \frac{(y_i - c_i)^2}{2\sigma_i^2} \right) \times \left(e^{-(Y_1 + c_1)^2/2\sigma_1^2} - e^{-(Y_1 - c_1)^2/2\sigma_1^2} \right). \quad (20)$$

The integrand (and hence the integral) is strictly positive for $c_1 < 0$ and negative for $c_1 > 0$. That is, for any fixed values of c_2, \dots, c_n , the function $P_{\mathbf{c}}$ has its only maximum at $c_1 = 0$. The same argument applies to each of the other c_j . Hence $P_{\mathbf{c}}$ has a global maximum at $\mathbf{c} = 0$.

[1] G. Hinshaw, D. N. Spergel, L. Verde, R. S. Hill, S. S. Meyer, C. Barnes, C. L. Bennett, M. Halpern, N. Jarosik, A. Kogut, et al., *Astrophys. J. Supp.* **148**, 135 (2003), arXiv:astro-ph/0302217.

[2] C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, D. N. Spergel, G. S. Tucker, et al., *Astrophys. J. Supp.* **148**, 1 (2003), arXiv:astro-ph/0302207.

[3] G. Hinshaw, M. R. Nolta, C. L. Bennett, R. Bean, O. Doré, M. R. Greason, M. Halpern, R. S. Hill, N. Jarosik, A. Kogut, et al., *Astrophys. J. Supp.* **170**, 288 (2007), arXiv:astro-ph/0603451.

[4] G. Hinshaw, J. L. Weiland, R. S. Hill, N. Odegard, D. Larson, C. L. Bennett, J. Dunkley, B. Gold, M. R. Greason, N. Jarosik, et al., ArXiv e-prints **803** (2008), 0803.0732.

[5] J. Dunkley, E. Komatsu, M. R. Nolta, D. N. Spergel, D. Larson, G. Hinshaw, L. Page, C. L. Bennett, B. Gold, N. Jarosik, et al., ArXiv e-prints **803** (2008), 0803.0586.

[6] E. Komatsu, J. Dunkley, M. R. Nolta, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. Limon, L. Page, et al., ArXiv e-prints **803** (2008), 0803.0547.

[7] C. J. Copi, D. Huterer, D. J. Schwarz, and G. D. Starkman, *Phys. Rev. D* **75**, 023507 (2007), arXiv:astro-ph/0605135.

[8] A. de Oliveira-Costa, M. Tegmark, M. Zaldarriaga, and A. Hamilton, *Phys. Rev. D* **69**, 063516 (2004), arXiv:astro-ph/0307282.

[9] D. J. Schwarz, G. D. Starkman, D. Huterer, and C. J. Copi, *Physical Review Letters* **93**, 221301 (2004), arXiv:astro-ph/0403353.

[10] C. J. Copi, D. Huterer, and G. D. Starkman, *Phys. Rev. D* **70**, 043515 (2004), arXiv:astro-ph/0310511.

[11] A. Hajian, ArXiv Astrophysics e-prints (2007), astro-ph/0702723.

[12] H. K. Eriksen, F. K. Hansen, A. J. Banday, K. M. Górski, and P. B. Lilje, *Astrophys. J.* **605**, 14 (2004).

[13] P. E. Freeman, C. R. Genovese, C. J. Miller, R. C. Nichol, and L. Wasserman, *Astrophys. J.* **638**, 1 (2006), arXiv:astro-ph/0510406.

[14] L. Campanelli, P. Cea, and L. Tedesco, *Physical Review Letters* **97**, 131302 (2006).

[15] J. D. Barrow, P. G. Ferreira, and J. Silk, *Physical Review Letters* **78**, 3610 (1997), arXiv:astro-ph/9701063.

[16] T. Ghosh, A. Hajian, and T. Souradeep, *Phys. Rev. D* **75**, 083007 (2007), arXiv:astro-ph/0604279.

[17] T. R. Jaffe, A. J. Banday, H. K. Eriksen, K. M. Górski, and F. K. Hansen, *Astron. Astrophys.* **460**, 393 (2006), arXiv:astro-ph/0606046.

[18] K. T. Inoue and J. Silk, *Astrophys. J.* **648**, 23 (2006), arXiv:astro-ph/0602478.

[19] K. T. Inoue and J. Silk, *Astrophys. J.* **664**, 650 (2007), arXiv:astro-ph/0612347.

[20] L. R. Abramo, L. J. Sodré, and C. A. Wuensche, *Phys. Rev. D* **74**, 083515 (2006), arXiv:astro-ph/0605269.

[21] A. Cooray and N. Seto, *Journal of Cosmology and Astro-*

Particle Physics **12**, 4 (2005), arXiv:astro-ph/0510137.

[22] P. C. Frisch, *Astrophys. J. Lett.* **632**, L143 (2005), arXiv:astro-ph/0506293.

[23] G. Efstathiou, M.N.R.A.S. **348**, 885 (2004), arXiv:astro-ph/0310207.

[24] I. Szapudi, S. Prunet, and S. Colombi, *Astrophys. J. Lett.* **561**, L11 (2001).

[25] A. Gruppuso, *Phys. Rev. D* **76**, 083010 (2007), arXiv:0705.2536.

[26] A. Gould, *Astrophys. J. Lett.* **403**, L51 (1993).

[27] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, et al., *Astrophys. J. Lett.* **396**, L1 (1992).

[28] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, *Astrophys. J.* **622**, 759 (2005), arXiv:astro-ph/0409513.

[29] D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, et al., *Astrophys. J. Supp.* **170**, 377 (2007), arXiv:astro-ph/0603449.

[30] D. N. Spergel, L. Verde, H. V. Peiris, E. Komatsu, M. R. Nolta, C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, et al., *Astrophys. J. Supp.* **148**, 175 (2003), arXiv:astro-ph/0302209.

[31] G. Efstathiou, M.N.R.A.S. **346**, L26 (2003), arXiv:astro-ph/0306431.

[32] E. Gaztañaga, J. Wagg, T. Multamäki, A. Montaña, and D. H. Hughes, M.N.R.A.S. **346**, 47 (2003), arXiv:astro-ph/0304178.

[33] J. M. Cline, P. Crotty, and J. Lesgourgues, *Journal of Cosmology and Astro-Particle Physics* **9**, 10 (2003), arXiv:astro-ph/0304558.

[34] C. R. Contaldi, M. Peloso, L. Kofman, and A. Linde, *Journal of Cosmology and Astro-Particle Physics* **7**, 2 (2003), arXiv:astro-ph/0303636.

[35] A. A. Starobinskij, *Soviet Journal of Experimental and Theoretical Physics Letters* **57**, 622 (1993), arXiv:gr-qc/9305019.

[36] D. Stevens, D. Scott, and J. Silk, *Physical Review Letters* **71**, 20 (1993).

[37] A. de Oliveira-Costa, G. F. Smoot, and A. A. Starobinsky, *Astrophys. J.* **468**, 457 (1996), arXiv:astro-ph/9510109.

[38] E. F. Bunn and D. Scott, M.N.R.A.S. **313**, 331 (2000), arXiv:astro-ph/9906044.

[39] J.-P. Luminet, J. R. Weeks, A. Riazuelo, R. Lehoucq, and J.-P. Uzan, *Nature (London)* **425**, 593 (2003), arXiv:astro-ph/0310253.

[40] N. J. Cornish, D. N. Spergel, G. D. Starkman, and E. Komatsu, *Physical Review Letters* **92**, 201302 (2004), arXiv:astro-ph/0310233.

[41] C. Gordon, W. Hu, D. Huterer, and T. Crawford, *Phys. Rev. D* **72**, 103002 (2005), arXiv:astro-ph/0509301.